Journal of Multivariate Analysis **78**, 62–82 (2001)
doi:10.1006/jmva.2000.1939, available online at http://www.idealibrary.com on IDEAL®

Inferences on a Normal Covariance Matrix and Generalized Variance with Monotone Missing Data

Jian Hao

St. Cloud State University

and

K. Krishnamoorthy[1]

University of Louisiana at Lafayette

Received June 14, 1999; published online February 5, 2001

The problems of testing a normal covariance matrix and an interval estimation of generalized variance when the data are missing from subsets of components are considered. The likelihood ratio test statistic for testing the covariance matrix is equal to a specified matrix, and its asymptotic null distribution is derived when the data matrix is of a monotone pattern. The validity of the asymptotic null distribution and power analysis are performed using simulation. The problem of testing the normal mean vector and a covariance matrix equal to a given vector and matrix is also addressed. Further, an approximate confidence interval for the generalized variance is given. Numerical studies show that the proposed interval estimation procedure is satisfactory even for small samples. The results are illustrated using simulated data. © 2001 Academic Press

AMS 1991 subject classifications: 62F25; 62H99.

Key words and phrases: generalized variance; likelihood ratio test; missing data; monotone patterns power; Satterthwaite approximation.

1. INTRODUCTION

The problem of missing data is an important applied problem, because missing values are commonly encountered in many practical situations. For instance, if some of the variables to be measured are too expensive, then experimenters may collect the complete data from a subset of the selected sample and collect data only on less expensive variables from the remaining individuals in the sample. In some situations, incomplete data arise because

[1] Corresponding author: K. Krishnamoorthy, Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA 70504-1010. Fax: (337) 482 4356. E-mail: krishna@louisiana.edu.

0047-259X/01 $35.00

of the accidental loss of data on some variables or variables are unobservable through design consideration. Two commonly used approaches to analyze incomplete data are the likelihood-based approach and the multiple imputation. The imputation method is to impute the missing data to form complete data and then use the standard methods available for the complete data analysis. For a good exposition of the imputation procedures and validity of imputation inferences in practice, we refer to Little and Rubin (1987) and Meng (1994). The results based on imputation methods are satisfactory even for nonnormal data and they are valid under the assumption that the data are missing at random (Rubin, 1976). However, inferences based on imputational method are in general valid only for large samples (Rubin and Schenker, 1986; Li *et al.*, 1991). If normality is assumed then the computational task involved in the imputation procedure can be avoided and simple test procedures for various patterns of data can be obtained. Furthermore, as shown in Krishnamoorthy and Pannala (1998, 1999), the results under the normality assumption are very satisfactory even for small samples although it requires the assumption that the data are missing completely at random (MCAR). MCAR means that missingness is completely independent of either the nature or the values of the variables in the data set under study. Practical situations where MCAR can be assumed are given at the beginning of this paragraph.

The problems treated in this article concern making inferences on normal covariance matrix Σ based on a monotone sample. A monotone sample can be described as follows. Consider a sample of N_1 individuals on which we are interested in p characteristics or variables. For the reasons given at the beginning of this section, we may not be able to observe all the variables from all the individuals in the sample. Suppose that we observe only p_1 variables from all N_1 individuals, $p_1 + p_2$ variables from a subset of N_2 individuals and so on; then the resulting sample data can be written in the following pattern known as a monotone pattern or triangular pattern.

$$\begin{array}{l} x_{11}, ..., x_{1N_k}, ..., x_{1N_2}, ..., x_{1N_1} \\ x_{21}, ..., x_{2N_k}, ..., x_{2N_2} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ x_{k1}, ..., x_{kN_k} \end{array} \tag{1.1}$$

We note that each x in the ith row represents $p_i \times 1$ vector observation, $i = 1, ..., k$, $p_1 + p_2 + \cdots + p_k = p$ and $N_1 > N_2 > \cdots > N_k$. In other words, $N_1 - N_k$ observations are missing on the last set of p_k components, $N_1 - N_{k-1}$ observations are missing on the last but first set of p_{k-1} components, and so on. Even though there are other missing patterns considered in the literature (for example, Lord, 1955), the monotone pattern is relatively easy to handle and is the most common in practice, and so we consider only monotone samples in this paper.

Several authors have considered the problem of making inferences about the normal mean vector μ and the covariance matrix Σ for monotone samples. Anderson (1957) provided a simple unified approach to derive the maximum likelihood estimators (MLEs) for μ and Σ and gave explicit expressions of the MLEs for the pattern of data (1.1) with $k = 2$. Following the method of Anderson, Bhargava (1962) derived likelihood ratio tests (LRTs) for many problems when the samples are of monotone pattern with $p_1 = \cdots = p_k = 1$. Since Bhargava's work, many authors have addressed the problems of confidence estimation and power analysis of the LRT about μ for some special cases; for example, see Mehta and Gurland (1969), Morrison (1973), and Naik (1975). Recently, Krishnamoorthy and Pannala (1998, 1999) developed a confidence region and some simple test procedures for μ when samples are of pattern (1.1) or Lord's (1955) pattern.

Although quite extensive studies are made on inferential procedures about the normal mean vector μ, only very limited results are available on the covariance matrix Σ. Regarding point estimation of Σ, Eaton (1970) developed a minimax estimator of Σ and showed that it is better than the MLE under an entropy loss function. Sharma and Krishnamoorthy (1985) derived an orthogonal invariant estimator which is better than Eaton's minimax estimator under the entropy loss. These results are for monotone samples. Bhargava (1962) derived the LRT statistic for testing the equality of several covariance matrices and an approximation to its null distribution.

In the following we summarize the results of this article.

In Section 2, we give some preliminary results. In Section 3, we consider the problem of testing the normal covariance matrix when the sample is of monotone pattern (1.1). We derive the LRT statistic using the conditional likelihood approach of Anderson (1957). Since the exact null distribution of the LRT statistic is difficult to derive, we approximate it by a constant times chi-square distribution using the Satterthwaite approximation. The validity of the approximate distribution is verified using simulation. Our numerical studies indicate that the results are satisfactory even for small samples. Power comparisons between the LRT based on incomplete data and the LRT based on complete data obtained by discarding the extra data are made to demonstrate the advantage of keeping the extra data; comparison studies indicate that the former is more powerful than the latter in all the cases considered. The results are extended to simultaneous testing hypotheses about the mean vector and covariance matrix in Section 4.

The problem of estimating the generalized variance $|\Sigma|$ is addressed in Section 5. We consider $|\hat{\Sigma}|$, where $\hat{\Sigma}$ is the MLE of Σ, as a point estimator of $|\Sigma|$. A traditional normal approximation (Section 5.1) and a new chi-square approximation (Section 5.2) to the distribution of $|\hat{\Sigma}|/|\Sigma|$ are derived. On the basis of the approximate distributions, we give confidence intervals for the generalized variance. We also showed that the chi-square approximation is good even for small samples whereas the normal approximation is not satisfactory even for moderately large samples. In Section 6, the results are illustrated using a simulated data set.

2. PRELIMINARIES

In the following we present some basic results in the notations of Krishnamoorthy and Pannala (1998). Let $X^{(l)}$ denote the submatrix of (1.1) formed by the first $p_1 + p_2 + \cdots + p_l$ rows and the first N_l columns, $l = 1, ..., k$. Let $\bar{x}^{(l)}$ and $S^{(l)}$ denote respectively the sample mean vector and the sums of squares and products matrix based on $X^{(l)}$, $l = 1, ..., k$. We shall give the results for the case of $k = 2$. The following results are well known.

$$\bar{x}^{(1)} = \bar{x}_{1,1} \sim N_{p_1}(\mu_1, \Sigma_{11}/N_1),$$

$$S^{(1)} = S_{11,1} \sim W_{p_1}(N_1 - 1, \Sigma_{11}),$$

$$\bar{x}^{(2)} = \begin{pmatrix} \bar{x}_{1,2} \\ \bar{x}_{2,2} \end{pmatrix} \sim N_p(\mu, \Sigma/N_2) \qquad \text{and}$$

$$S^{(2)} = \begin{pmatrix} S_{11,2} & S_{12,2} \\ S_{21,2} & S_{22,2} \end{pmatrix} \sim W_p(N_2 - 1, \Sigma).$$

Let $\mu_{2.1} = \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1$ and $\Sigma_{2.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$. The maximum likelihood estimators of the parameters are (Anderson, 1957) given by

$$\hat{\mu}_1 = \bar{x}_{1,1}, \qquad \hat{\mu}_{2.1} = \bar{x}_{2,2} - S_{21,2}S_{11,2}^{-1}\bar{x}_{1,2},$$

$$\hat{\Sigma}_{11} = S_{11,1}/N_1, \qquad \text{and} \qquad \hat{\Sigma}_{2.1} = S_{2.1,2}/N_2$$

$$= (S_{22,2} - S_{21,2}S_{11,2}^{-1}S_{12,2})/N_2. \qquad (2.1)$$

These are the MLEs that we will use to derive the likelihood ratio tests for testing Σ and for testing μ and Σ simultaneously.

The following lemmas are needed in the remainder of the paper.

LEMMA 2.1. *Let $S \sim W_p(n, I_p)$. Write S as*

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

so that S_{11} *is of order* $p_1 \times p_1$ *and* S_{22} *is of order* $p_2 \times p_2$. *Then,* S_{11}, $S_{2.1.}$, *and* $S_{12}S_{22}^{-1}S_{21}$ *are statistically independent with*

(i) $S_{11} \sim W_{p_1}(n, I_p)$,
(ii) $S_{2.1} \sim W_{p_2}(n - p_1, I_{p_2})$,
(iii) $\mathrm{tr}(S_{12}S_{22}^{-1}S_{21}) \sim \chi^2_{p_1 p_2}$.

Proof. See Siotani *et al.* (1985, Theorem 2.4.1, p. 70).

A proof of the following lemma can be found, for example, in Muirhead (1982, p. 100).

LEMMA 2.2. *Let* $S \sim W_p(n, \Lambda)$, *where* Λ *is a positive definite matrix. Then,*

$$\frac{|S|}{|\Lambda|} \sim \prod_{i=1}^{p} \chi^2_{n-i+1},$$

where all the chi-square variates are independent.

3. LIKELIHOOD RATIO TEST FOR Σ

We derive the LRT statistic for testing $H_0: \Sigma = \Sigma_0$ against $H_a: \Sigma \neq \Sigma_0$, where Σ_0 is a specified positive definite matrix when the data set is of monotone pattern (1.1) with $k = 2$. It is easy to see that the testing problem is invariant under the transformations

$$\begin{pmatrix} x_{11}, \ldots, x_{1N_2} \\ x_{21}, \ldots, x_{2N_2} \end{pmatrix} \to A \begin{pmatrix} x_{11}, \ldots, x_{1N_2} \\ x_{21}, \ldots, x_{2N_2} \end{pmatrix}$$

and

$$(x_{1N_2+1}, \ldots, x_{1N_1}) \to A_{11}(x_{1N_2+1}, \ldots, x_{1N_1}),$$

where A is a $p \times p$ nonsingular matrix with the submatrix $A_{12} = 0$, and A_{11} is the (1,1) submatrix of A with order $p_1 \times p_1$. Therefore, without loss of generality, we can assume that $\Sigma_0 = I_p$, the identity matrix of order p.

3.1. *LRT Statistic*

Let $n_p(x \mid \mu, \Sigma)$ denote the probability density function of a p-variate normal random vector with mean μ and covariance matrix Σ. Writing the joint density as the product of the marginal and conditional density functions, we can express the likelihood function as

$$L(\mu, \Sigma) = \prod_{j=1}^{N_1} n_{p_1}(x_{1j} \mid \mu_1, \Sigma_{11}) \prod_{j=1}^{N_2} n_{p_2}(x_{2j} \mid \mu_{2.1} + \Sigma_{21}\Sigma_{11}^{-1}x_{1j}, \Sigma_{2.1}), \tag{3.1}$$

where $\mu_{2.1} = \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1$ and $\Sigma_{2.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$. Using the MLEs given in (2.1), it can be shown that the LRT statistic is

$$\frac{\sup_\mu L(\mu, I_p)}{\sup_{\mu, \Sigma} L(\mu, \Sigma)} = e^{(N_1 p_1 + N_2 p_2)/2} \, |S_{11,1}/N_1|^{N_1/2} \, |\hat{\Sigma}_{2.1}|^{N_2/2} \times \exp\left\{ -\frac{1}{2} [\mathrm{tr}(S_{11,1} + S_{22,2})] \right\}. \tag{3.2}$$

It is well known that the LRT for testing $\Sigma = \Sigma_0$ is biased even when no data are missing. Further, by replacing the sample size by the degrees of freedom, an unbiased test can be obtained (Das Gupta, 1969). In view of this fact, we modify the LRT statistic (3.2) by replacing N_1 by $n_1 = N_1 - 1$ and N_2 by $n_2 = N_2 - p_1 - 1$. The reason for choosing n_1 and n_2 is that $S_{11,1} \sim W_{p_1}(n_1, I_{p_1})$ and $N_2\hat{\Sigma}_{2.1} \sim W_{p_2}(n_2, I_{p_2})$. The modified LRT statistic is given by

$$\Lambda = \left\{ \left(\frac{e}{n_1}\right)^{n_1 p_1/2} |S_{11,1}|^{n_1/2} \exp\left(-\frac{1}{2}\mathrm{tr}(S_{11,1})\right) \right\} \times \left\{ \left(\frac{e}{n_2}\right)^{n_2 p_2/2} |S_{2.1,2}|^{n_2/2} \exp\left(-\frac{1}{2}\mathrm{tr}(S_{22,2})\right) \right\}.$$

Using the relation that $S_{22,2} = S_{2.1,2} + S_{21,2}S_{11,2}^{-1}S_{12,2}$, we can write Λ as

$$\Lambda = \Lambda_1 \Lambda_2 \exp\{ -\tfrac{1}{2}\mathrm{tr}(S_{21,2}S_{11,2}^{-1}S_{12,2})\}, \tag{3.3}$$

where

$$\Lambda_1 = \left(\frac{e}{n_1}\right)^{n_1 p_1/2} |S_{11,1}|^{n_1/2} \exp\left(-\frac{1}{2}\mathrm{tr}(S_{11,1})\right)$$

and

$$\Lambda_2 = \left(\frac{e}{n_2}\right)^{n_2 p_2/2} |S_{2.1,2}|^{n_2/2} \exp\left(-\frac{1}{2}\mathrm{tr}(S_{2.1,2})\right).$$

3.2. *An Asymptotic Null Distribution of the Modified LRT Statistic*

The exact null distribution of Λ is difficult to obtain. Therefore, we resorted to finding an asymptotic null distribution of Λ. It follows from Lemma 2.1 that Λ_1, Λ_2, and $\mathrm{tr}(S_{21,2}S_{11,2}^{-1}S_{12,2})$ are statistically independent with $\mathrm{tr}(S_{21,2}S_{11,2}^{-1}S_{12,2}) \sim \chi^2_{p_1 p_2}$. Further, it can be deduced from Muirhead (1982, p. 359) that asymptotically

$$-2 \ln \Lambda_i \sim \chi^2_{f_i}/\rho_i,$$

where

$$f_i = p_i(p_i+1)/2,$$

$$\rho_i = 1 - \frac{2p_i^2 + 3p_i - 1}{6n_i(p_i+1)}, \qquad i = 1, 2.$$

Therefore, it follows from (3.3) that asymptotically

$$\begin{aligned} -2 \ln \Lambda &= -2 \ln \Lambda_1 - 2 \ln \Lambda_2 + \mathrm{tr}(S_{21,2}S_{11,2}^{-1}S_{12,2}) \\ &\sim \frac{1}{\rho_1}\chi^2_{f_1} + \frac{1}{\rho_2}\chi^2_{f_2} + \chi^2_{p_1 p_2}. \end{aligned} \tag{3.4}$$

Thus, $-2 \ln \Lambda$ is asymptotically distributed as a linear combination of three independent chi-square random variables. Again, it is not easy to find the exact distribution of

$$W = \frac{1}{\rho_1}\chi^2_{f_1} + \frac{1}{\rho_2}\chi^2_{f_2} + \chi^2_{p_1 p_2}.$$

So we approximate the distribution of W by using the classical Satterthwaite's (1946) approximation. That is, we approximate the distribution of W by the distribution of $a\chi^2_b$, where the positive constants a and b are estimated as the solutions of the equations $E(W) = E(a\chi^2_b)$ and $E(W^2) = E(a\chi^2_b)^2$. Toward this, we note that

$$E(W) = \frac{1}{\rho_1} f_1 + \frac{1}{\rho_2} f_2 + p_1 p_2 = M_1 \qquad \text{(say)},$$

and

$$E(W^2) = M_1^2 + \frac{2}{\rho_1^2} f_1 + \frac{2}{\rho_2^2} f_2 + 2p_1 p_2 = M_2 \qquad \text{(say)}.$$

TABLE I

The Simulated Sizes of the Modified LRT

		$(p_1=2, p_2=1)$		$(p_1=2, p_2=2)$	
N_1	N_2	$\alpha=0.05$	$\alpha=0.01$	$\alpha=0.05$	$\alpha=0.01$
20	10	0.0501	0.0103	0.0505	0.0101
30	15	0.0509	0.0102	0.0504	0.0103
30	20	0.0505	0.0104	0.0504	0.0102
40	30	0.0506	0.0101	0.0501	0.0100
60	40	0.0493	0.0103	0.0502	0.0097
80	50	0.0501	0.0101	0.0500	0.0103

Equating M_1 and M_2 respectively to the first and the second moments of $a\chi_b^2$, and solving the resulting equations for a and b, we see that

$$a = \frac{M_2 - M_1^2}{2M_1} \qquad \text{and} \qquad b = \frac{M_1}{a}.$$

Thus, under $H_0\colon \Sigma = I_p$,

$$-2\ln \Lambda \;\dot{\sim}\; a\chi_b^2. \tag{3.5}$$

The notation $\dot{\sim}$ should be interpreted as "approximately distributed as." Therefore, the size of the test that rejects $H_0\colon \Sigma = \Sigma_0$ whenever

$$-2\ln \Lambda > a\chi_b^2(1-\alpha), \tag{3.6}$$

where $a\chi_b^2(1-\alpha)$ denotes the $100(1-\alpha)$th percentile of the $a\chi_b^2$ random variable, is approximately equal to α.

To understand the validity of the asymptotic null distribution in (3.5), we simulated the sizes of the test (using 100,000 runs) at levels $\alpha = 0.05$ and 0.01. The IMSL subroutine RNMVN is used to generate multivariate normal variates. For given N_1, N_2, p_1, p_2, and α, the size of the test is estimated by the proportion of the times $-2\ln \Lambda$ exceeds the $100(1-\alpha)$th percentile of $a\chi_b^2$. The simulation results are presented in Table I. It is clear from Table I that the approximation is very satisfactory even for small values of N_1 and N_2.

3.3. *Power Studies*

To understand the nature of the powers of the modified LRT, and the advantage of keeping additional data, we estimated the powers of the modified LRT and the modified LRT based on "partially complete data" (the data obtained after deleting the additional $N_1 - N_2$ observations on

TABLE II

Powers of the Modified LRT Based on Incomplete Data and on Partially Complete Data (in Parentheses) $p_1=2$, $p_2=1$; H_0: $\Sigma=I_p$

N_1	N_2	$\alpha=0.05$	$\alpha=0.01$	$\alpha=0.05$	$\alpha=0.01$
		H_a: $\Sigma=\text{diag}(0.8, 0.6, 0.5)$		H_a: $\Sigma=\text{diag}(0.5, 0.4, 0.2)$	
12	7	0.1180(0.0933)	0.0304(0.0225)	0.5594(0.2455)	0.2731(0.0775)
15	10	0.1602(0.1150)	0.0460(0.0297)	0.7527(0.4150)	0.4723(0.1679)
20	10	0.2091(0.0999)	0.0661(0.0245)	0.8868(0.3776)	0.6810(0.1464)
25	15	0.2915(0.1481)	0.1088(0.0425)	0.9722(0.6643)	0.8783(0.3724)
30	20	0.3844(0.2103)	0.1643(0.0682)	0.9951(0.8588)	0.9667(0.6281)
40	30	0.5649(0.3913)	0.3029(0.1696)	0.9999(0.9943)	0.9989(0.9619)
		H_a: $\Sigma=\begin{pmatrix} 2 & 0.4 & 0 \\ 0.4 & 1.5 & 0 \\ 0 & 0 & 1.3 \end{pmatrix}$		H_a: $\Sigma=\begin{pmatrix} 2 & 0.4 & -0.3 \\ 0.4 & 1.5 & 0 \\ -0.3 & 0 & 1.3 \end{pmatrix}$	
12	7	0.3226(0.2213)	0.1603(0.0901)	0.3353(0.2213)	0.1696(0.0901)
15	10	0.4002(0.3036)	0.2164(0.1445)	0.4111(0.3066)	0.2262(0.1445)
20	10	0.5043(0.3666)	0.3030(0.1878)	0.5030(0.3666)	0.3036(0.1878)
25	15	0.6091(0.5016)	0.4021(0.2997)	0.6075(0.5016)	0.4011(0.2997)
30	20	0.6882(0.6065)	0.4893(0.4029)	0.6937(0.6065)	0.4983(0.4029)
40	30	0.8184(0.7682)	0.6538(0.5882)	0.8165(0.7672)	0.6518(0.5872)

the first p_1 components) using simulation. The powers of the modified LRT based on partially complete data are given in parentheses. The powers are computed when H_0: $\Sigma=I_p$ and presented in Table II.

We observe from Table II that the powers of the tests are increasing as sample sizes increase; they are also increasing as Σ moves away from the specified matrix I_p in H_0. Further, the powers of the modified LRT based on incomplete data are always larger than the corresponding powers of the modified LRT based on partially complete data. This demonstrates the advantage of keeping the extra data available on the first p_1 components.

3.4. *Generalization*

The LRT procedures discussed in the earlier sections can be extended to the monotone pattern (1.1) with $k\geqslant 3$ in a relatively easy manner. For convenience, let us first illustrate the case $k=3$ and the null hypothesis H_0: $\Sigma=I_p$. The modified LRT statistic is given by

$$\Lambda=\Lambda_1\Lambda_2\Lambda_3\exp\{-\tfrac{1}{2}\operatorname{tr}(S_{21,2}S_{11,2}^{-1}S_{12,2})\}$$
$$\times\exp\left\{-\tfrac{1}{2}\operatorname{tr}(S_{31,3}, S_{32,3})\begin{pmatrix} S_{11,3} & S_{12,3} \\ S_{21,3} & S_{22,3} \end{pmatrix}^{-1}\begin{pmatrix} S_{13,3} \\ S_{23,3} \end{pmatrix}\right\}, \tag{3.7}$$

where Λ_1, Λ_2, n_1, n_2 are as defined in Section 3.2,

$$\Lambda_3 = \left(\frac{e}{n_3}\right)^{p_3 n_3/2} |S_{3.21,3}|^{n_3/2} \exp\left\{-\frac{1}{2} \operatorname{tr} S_{3.21,3}\right\},$$

and $n_3 = N_3 - (p_1 + p_2) - 1$.

It follows from a straightforward extension of Lemma 2.1 that all five terms in (3.7) are independent with

$$\operatorname{tr}\left[(S_{31,3}, S_{32,3}) \begin{pmatrix} S_{11,3} & S_{12,3} \\ S_{21,3} & S_{22,3} \end{pmatrix}^{-1} \begin{pmatrix} S_{13,3} \\ S_{23,3} \end{pmatrix}\right] \sim \chi^2_{(p_1+p_2)\,p_3},$$

and asymptotically

$$-2 \ln \Lambda_3 \sim \frac{1}{\rho_3} \chi^2_{f_3},$$

where $f_3 = p_3(p_3+1)/2$ and $\rho_3 = 1 - (2p_3^2 + 3p_3 - 1)/(6n_3(p_3+1))$. Thus, asymptotically

$$-2 \ln \Lambda \mathrel{\dot\sim} \frac{1}{\rho_1} \chi^2_{f_1} + \frac{1}{\rho_2} \chi^2_{f_2} + \frac{1}{\rho_3} \chi^2_{f_3} + \chi^2_{p_1 p_2} + \chi^2_{(p_1+p_2)\,p_3},$$

and hence its null distribution can be approximated by the distribution of a constant times chi-square random variable as in Section 3.2.

The expression of Λ for a general k can be obtained similarly using the MLEs (see, for example, Jinadasa and Tracy, 1992) of μ and Σ. Following the notations defined at the beginning of Section 2, we can partition the sample summary statistics as

$$\bar{x}^{(l)} = \begin{pmatrix} \bar{x}_{1,l} \\ \bar{x}_{2,l} \\ \vdots \\ \bar{x}_{l,l} \end{pmatrix} \quad \text{and} \quad S^{(l)} = \begin{pmatrix} S_{11,l} & \cdots & S_{1l,l} \\ \vdots & & \vdots \\ S_{l1,l} & \cdots & S_{ll,l} \end{pmatrix}, \qquad l = 1, ..., k.$$

Further, define

$$(B_{l1}, ..., B_{l(l-1)}) = (S_{l1,l}, ..., S_{l(l-1),l}) \begin{pmatrix} S_{11,l} & \cdots & S_{1(l-1),l} \\ \vdots & & \vdots \\ S_{(l-1)1,l} & \cdots & S_{(l-1)(l-1),l} \end{pmatrix}^{-1},$$

$$l = 2, ..., k. \tag{3.8}$$

Using these notations, the MLEs can be expressed as

$$\hat{\mu}_1 = \bar{x}^{(1)} = \bar{x}_{1,1}, \qquad \hat{\mu}_l = \bar{x}_{l,l} - \sum_{j=1}^{l-1} B_{lj}(\bar{x}_{j,l} - \hat{\mu}_j), \qquad \hat{\Sigma}_{11} = S^{(1)}/N_1;$$

and

$$N_l \hat{\Sigma}_{l\cdot(l-1)\cdots 1} = S_{l\cdot(l-1)\cdots 1,l} = S_{ll,l} - \sum_{j=1}^{l-1} B_{lj} S_{jl,l}, \qquad l = 2, ..., k. \tag{3.9}$$

In terms of these notations, the LRT statistic can be expressed as

$$e^{(1/2)\sum_{i=1}^{k} N_i p_i} \, |S_{11,1}/N_1|^{(1/2)N_1} \cdot |\hat{\Sigma}_{2.1}|^{(1/2)N_2} \cdots$$
$$|\hat{\Sigma}_{k\cdot(k-1)\cdots 21}|^{(1/2)N_k} \exp\left\{ -\tfrac{1}{2} \sum_{i=1}^{k} \operatorname{tr}(S_{ii,i}) \right\},$$

where $N_l \hat{\Sigma}_{l\cdot(l-1)\cdots 21}$ follows a Wishart distribution, $W_{p_l}(N_l - \sum_{i=1}^{l-1} p_i - 1, \Sigma_{l\cdot(l-1)\cdots 21})$, $l = 2, ..., k$. Further, letting $n_i = N_i - (p_1 + \cdots + p_{i-1}) - 1$, $i = 2, ..., k$, the modified LRT statistic can be written as

$$\Lambda = \left(\prod_{i=1}^{k} \Lambda_i \right) \prod_{i=2}^{k} \left\{ \exp\left(-\tfrac{1}{2} \operatorname{tr} \sum_{j=1}^{i-1} B_{ij} S_{ji,i} \right) \right\},$$

where

$$\Lambda_i = \left(\frac{e}{n_i} \right)^{n_i p_i/2} |S_{i\cdot(i-1)\cdots 21,i}|^{n_i/2} \exp\{ -\tfrac{1}{2} \operatorname{tr} S_{i\cdot(i-1)\cdots 21,i} \}, \quad i = 2, ..., k.$$

It follows again from a generalization of Lemma 2.1 that all the terms involved in Λ are independent with

$$-2 \ln \Lambda_i \sim \frac{1}{\rho_i} \chi^2_{f_i}, \qquad \text{asymptotically for} \quad i = 1, ..., k,$$

and

$$\operatorname{tr}\left(\sum_{j=1}^{i-1} B_{ij} S_{ji,i} \right) \sim \chi^2_{(p_1 + \cdots + p_{i-1})\, p_i}, \qquad i = 2, ..., k.$$

Therefore, asymptotically

$$-2 \ln \Lambda \sim \left(\sum_{i=1}^{k} \frac{1}{\rho_i} \chi^2_{f_i} \right) + \chi^2_{p_1 p_2} + \chi^2_{(p_1+p_2)\, p_3} + \cdots + \chi^2_{(p_1 + \cdots + p_{k-1})\, p_k}, \tag{3.10}$$

where

$$f_i = p_i(p_i+1)/2 \qquad \text{and} \qquad \rho = 1 - \frac{2p_i^2+3p_i-1}{6n_i(p_i+1)}, \qquad i=1, \ldots, k.$$

Since (3.10) is a linear combination of independent chi-square random variables, its distribution can be approximated by the distribution of a constant times chi-square random variable as in Section 3.2.

4. TESTING HYPOTHESES THAT A MEAN VECTOR AND A COVARIANCE MATRIX ARE EQUAL TO A GIVEN VECTOR AND MATRIX

We shall derive the LRT statistic for testing

$$H_0\colon \mu=\mu_0, \Sigma=\Sigma_0 \text{ vs } H_a\colon H_0 \text{ is not true.}$$

As in Section 3, without loss generality, we can assume that $\mu=0$ and $\Sigma=I_p$. For the case $k=2$, we derived the LRT statistic using the approach of Section 3.1, which is given by

$$\begin{aligned}\mathrm{LRTS} = e^{1/2(N_1p_1+N_2p_2)} &\left|\frac{1}{N_1} S_{11,1}\right|^{(1/2)N_1} |\hat{\Sigma}_{2.1}|^{(1/2)N_2} \\ &\times \exp\left\{-\frac{1}{2}\operatorname{tr}(S_{11,1})\right\} \exp\left\{-\frac{1}{2}\operatorname{tr}(S_{22,2})\right\} \\ &\times \exp\left\{-\frac{1}{2}N_1\bar{x}'_{1,1}\bar{x}_{1,1}\right\} \exp\left\{-\frac{1}{2}N_2\bar{x}'_{2,2}\bar{x}_{2,2}\right\}.\end{aligned}$$

Again for the same reason given in Section 3.1, we modify the LRT statistic by replacing N_1 by $n_1=N_1-1$ and N_2 by $n_2=N_2-p_1-1$ except for the last two terms. After modification and rearranging some terms, we can express the LRT statistic as

$$\delta = \Lambda \exp\{-\tfrac{1}{2}(N_1\bar{x}'_{1,1}\bar{x}_{1,1} + N_2\bar{x}'_{2,2}\bar{x}_{2,2})\}, \tag{4.1}$$

where Λ is the LRT statistic given in (3.3). Note that, under $H_0\colon \mu=\mu_0$, $\Sigma=I_p$, $N_1\bar{x}'_{1,1}\bar{x}_{1,1}$ and $N_2\bar{x}'_{2,2}\bar{x}_{2,2}$ are independent with $N_i\bar{x}'_{i,i}\bar{x}_{i,i} \sim \chi^2_{p_i}$, $i=1,2$. Therefore, $N_1\bar{x}'_{1,1}\bar{x}_{1,1} + N_2\bar{x}'_{2,2}\bar{x}_{2,2} \sim \chi^2_{p_1+p_2}$. Thus, it can be shown along the lines of Section 3.2 that asymptotically

$$-2\ln\delta \sim \frac{1}{\rho_1}\chi^2_{f_1} + \frac{1}{\rho_2}\chi^2_{f_2} + \chi^2_{p_1+p_2+p_1p_2},$$

where ρ_i's and f_i's are given in Section 3.2. The distribution of $\frac{1}{\rho_1}\chi^2_{f_1} + \frac{1}{\rho_2}\chi^2_{f_2} + \chi^2_{p_1+p_2+p_1p_2}$ can be approximated by the distribution of $c\chi^2_d$. Our numerical studies (not reported here; see Hao, 1999) indicated that this approximation is also satisfactory even for small sample sizes.

5. CONFIDENCE INTERVALS FOR THE GENERALIZED VARIANCE

In this section, we are interested in making inferences about the generalized variance $|\Sigma|$ with missing data of monotone pattern (1.1). We consider $|\hat{\Sigma}|$, where $\hat{\Sigma}$ is the MLE of Σ, as a point estimator of $|\Sigma|$. We give two confidence intervals for the generalized variance, one is based on the asymptotic normal distribution of $|\hat{\Sigma}|$ and the other is based on a chi-square approximation to the distribution of $(|\hat{\Sigma}|/|\Sigma|)^{1/p}$.

The determinant of $\hat{\Sigma}$ is

$$|\hat{\Sigma}| = |\hat{\Sigma}_{11}| \prod_{l=2}^{k} |\hat{\Sigma}_{l\cdot(l-1)\cdots 1}|, \tag{5.1}$$

where $\hat{\Sigma}_{11}$ and $\hat{\Sigma}_{l\cdot(l-1)\cdots 1}$ are given in (3.9). Further, it is well known that all k terms on the right hand side of (5.1) are independent with

$$N_l\hat{\Sigma}_{l\cdot(l-1)\cdots 1} \sim W_{p_l}(N_l - q_l - 1, \Sigma_{l\cdot(l-1)\cdots 1}), \qquad l = 1, ..., k, \tag{5.2}$$

where $q_1 = 0$, $q_i = p_1 + \cdots + p_{i-1}$, $i = 2, ..., k$, $\Sigma_{1.0} = \Sigma_{11}$, and $\hat{\Sigma}_{1.0} = \hat{\Sigma}_{11}$.

5.1. *Confidence Interval for $|\Sigma|$ Based on the Asymptotic Normal Distribution*

We will use the following asymptotic distribution of $|\hat{\Sigma}|$ to construct a confidence interval of $|\hat{\Sigma}|$.

THEOREM 5.1. *The asymptotic distribution of* $\ln|\hat{\Sigma}| - \ln|\Sigma|$ *is*

$$N\left(-\sum_{j=1}^{k} p_j \ln\frac{N_j}{N_j - q_j - 1},\ 2\sum_{j=1}^{k}\frac{p_j}{N_j - q_j - 1}\right). \tag{5.3}$$

Proof. For $i = 1, ..., k$, it follows from (5.2) that

$$\frac{N_i\hat{\Sigma}_{i\cdot(i-1)\cdots 1}}{N_i - q_i - 1} \sim W_{p_i}\left(N_i - q_i - 1, \frac{1}{N_i - q_i - 1}\Sigma_{i\cdot(i-1)\cdots 1}\right).$$

Therefore, by Theorem 3.2.16 of Muirhead (1982, p. 102),

$$\sqrt{\frac{N_i-q_i-1}{2p_i}}\left[\ln\left(\frac{N_i}{N_i-p_i-1}\right)^{p_i}+\ln\frac{|\hat{\Sigma}_{i\cdot(i-1)\cdots 1}|}{|\Sigma_{i\cdot(i-1)\cdots 1}|}\right]$$

$$\sim N(0, 1) \text{ for large } N_i, \qquad i=1, ..., k,$$

which implies that

$$\ln\frac{|\hat{\Sigma}_{i\cdot(i-1)\cdots 1}|}{|\Sigma_{i\cdot(i-1)\cdots 1}|}\sim N\left(-p_i\ln\frac{N_i}{N_i-q_i-1}, \frac{2p_i}{N_i-q_i-1}\right).$$

Using the relation $\ln|\hat{\Sigma}|-\ln|\Sigma|=\sum_{i=1}^{k}\ln(|\hat{\Sigma}_{i\cdot(i-1)\cdots 1}|/|\Sigma_{i\cdot(i-1)\cdots 1}|)$ and the fact that $\hat{\Sigma}_{i\cdot(i-1)\cdots 1}$, $i=1, ..., k$, are independent, we complete the proof.

Let $m=-\sum_{j=1}^{k}p_j\ln N_j/(N_j-q_j-1)$, $v=\sum_{j=1}^{k}2p_j/(N_j-q_j-1)$, and z_β denote the 100βth percentile of the standard normal distribution. For large samples, a $100(1-\alpha)\%$ confidence interval for the generalized variance $|\Sigma|$ based on (5.3) is given by

$$(|\hat{\Sigma}|/\exp(m+z_{1-\alpha/2}\sqrt{v}), |\hat{\Sigma}|/\exp(m-z_{1-\alpha/2}\sqrt{v})). \tag{5.4}$$

Our preliminary numerical studies indicated that the normal approximation is not satisfactory even for moderately large samples and so we give a chi-square approximation to the distribution of $|\hat{\Sigma}|/|\Sigma|$ in the following section.

5.2. *Confidence Interval for $|\Sigma|$ Based on a Chi-square Approximation*

In some special cases, the product of k independent chi-square random variables is distributed as a constant times the kth power of a chi-square random variable (see Anderson, 1984, p. 264). Further, Hoel (1937) suggested a constant times gamma distribution as an approximation to the distribution of the pth root of the sample generalized variance. These results indicate that the distribution of $(|\hat{\Sigma}|/|\Sigma|)^{1/p}$ can be approximated by a constant times chi-square distribution in an incomplete data setup as well. We approximate the distribution of $(|\hat{\Sigma}|/|\Sigma|)^{1/p}$ by the distribution of $a\chi_b^2$, where a and b are unknown positive parameters which can be estimated by the method of moments. Toward this, we need to find the moments of $|\hat{\Sigma}|$.

It follows from (5.2) and Lemma 2.2 that

$$\frac{|\hat{\Sigma}_{l\cdot(l-1)\cdots 1}|}{|\Sigma_{l\cdot(l-1)\cdots 1}|}\sim\frac{1}{N_l^{p_l}}\prod_{i=1}^{p_l}\chi^2_{N_l-q_l-i}, \qquad l=1, ..., k, \tag{5.5}$$

where all the chi-square variates are independent. Using this result, and the rth moment of a χ_a^2 random variable,

$$E(\chi_a^2)^r = \frac{2^r \Gamma(a/2+r)}{\Gamma(a/2)},$$

we get

$$E(|\hat{\Sigma}_{l\cdot(l-1)\cdots 1}|)^r = \frac{|\Sigma_{l\cdot(l-1)\cdots 1}|^r \, 2^{rp_l}}{N_l^{rp_l}} \prod_{i=1}^{p_l} \times \frac{\Gamma[(N_l - q_l - i)/2 + r]}{\Gamma[(N_l - q_l - i)/2]}, \qquad l = 1, ..., k. \tag{5.6}$$

By (5.1) and (5.2) and noticing that $|\Sigma| = \prod_{l=1}^{k} |\Sigma_{l\cdot(l-1)\cdots 1}|$, we get

$$E(|\hat{\Sigma}|)^r = |\Sigma|^r \, 2^{rp} \prod_{l=1}^{k} \left(\frac{1}{N_l^{rp_l}} \prod_{i=1}^{p_l} \frac{\Gamma[(N_l - q_l - i)/2 + r]}{\Gamma[(N_l - q_l - i)/2]} \right), \tag{5.7}$$

where $q_i = p_1 + \cdots + p_{i-1}$, $i = 2, ..., k$, $q_1 = 0$, and $p = p_1 + \cdots + p_k$.
Thus, from (5.7) we have

$$\begin{aligned} M_r &= E\left(\frac{|\hat{\Sigma}|}{|\Sigma|}\right)^{r/p} \\ &= 2^r \prod_{l=1}^{k} \prod_{i=1}^{p_l} \frac{\Gamma[(N_l - q_l - i)/2 + r/p]}{\Gamma[(N_l - q_l - i)/2]}, \qquad r = 1, 2. \end{aligned} \tag{5.8}$$

Equating M_1 and M_2 respectively to the first and second moment of $a\chi_b^2$, and solving the resulting equations for a and b, we get

$$a = \frac{M_2 - M_1^2}{2M_1} \qquad \text{and} \qquad b = M_1/a. \tag{5.9}$$

Thus,

$$\left(\prod_{l=1}^{k} N_l^{p_l}\right)^{1/p} \left(\frac{|\hat{\Sigma}|}{|\Sigma|}\right)^{1/p} = \left(\prod_{l=1}^{k} \prod_{i=1}^{p_l} \chi^2_{N_l - q_l - i}\right)^{1/p} \mathrel{\dot{\sim}} a\chi_b^2, \tag{5.10}$$

where a and b are given in (5.9). Furthermore, for given $0 < \beta < 1$, let c_β denote the 100βth percentile of $|\hat{\Sigma}|/|\Sigma|$. Then,

$$c_\beta \doteq \frac{[a\chi_b^2(\beta)]^p}{\prod_{l=1}^{k} N_l^{p_l}}, \tag{5.11}$$

TABLE III

Approximate Percentiles and the Simulated Cumulative Probabilities (Given in Parentheses) $p_1 = p_2 = p_3 = 1$

N_1	N_2	N_3	χ^2	Normal	χ^2	Normal	χ^2	Normal
			$\alpha = 0.10$		$\alpha = 0.05$		$\alpha = 0.01$	
12	8	5	0.0187	0.0568	0.0099	0.0363	0.0027	0.0157
			(0.1097)	(0.2745)	(0.0617)	(0.1939)	(0.0172)	(0.0943)
16	13	10	0.1403	0.2056	0.1001	0.1552	0.0515	0.915
			(0.1006)	(0.1965)	(0.0511)	(0.1213)	(0.0112)	(0.0422)
25	19	15	0.2524	0.3159	0.1964	0.2535	0.1204	0.1677
			(0.0998)	(0.1702)	(0.0503)	(0.1009)	(0.0106)	(0.0312)
40	30	25	0.3905	0.4429	0.3247	0.3744	0.2273	0.2733
			(0.1014)	(0.1524)	(0.0507)	(0.0874)	(0.0103)	(0.0248)
55	50	40	0.5065	0.5466	0.4395	0.4788	0.3346	0.3735
			(0.1005)	(0.1391)	(0.0504)	(0.0772)	(0.0101)	(0.0202)
65	60	55	0.5606	0.5950	0.4949	0.5292	0.3897	0.4247
			(0.1004)	(0.1341)	(0.0497)	(0.0737)	(0.0100)	(0.0185)
120	90	80	0.6480	0.6723	0.5878	0.5145	0.4884	0.5145
			(0.1001)	(0.1273)	(0.0498)	(0.0680)	(0.0101)	(0.0164)
			$\alpha = 0.90$		$\alpha = 0.95$		$\alpha = 0.99$	
12	8	5	0.6758	0.13318	0.10163	2.0828	2.0759	4.8189
			(0.8991)	(0.9726)	(0.9513)	(0.9913)	(0.9913)	(0.9997)
16	13	10	1.1267	1.4995	1.4603	1.9872	2.3268	3.3703
			(0.8994)	(0.9546)	(0.9506)	(0.9825)	(0.9907)	(0.9983)
25	19	15	1.2500	1.4946	1.5345	1.8629	2.2258	2.8161
			(0.9005)	(0.9454)	(0.9507)	(0.9776)	(0.9902)	(0.9972)
40	30	25	1.3021	1.4480	1.5256	1.7127	2.0383	2.3468
			(0.8994)	(0.9360)	(0.9497)	(0.9725)	(0.9906)	(0.9967)
55	50	40	1.3003	1.3907	1.4749	1.5875	1.8595	2.0350
			(0.9003)	(0.9298)	(0.9498)	(0.9686)	(0.9904)	(0.9955)
65	60	55	1.2910	1.3609	1.4443	1.5303	1.7761	1.9068
			(0.9004)	(0.9269)	(0.9497)	(0.9674)	(0.9905)	(0.9952)
120	90	80	1.2527	1.2956	1.3702	1.4218	1.6173	1.6928
			(0.9000)	(0.9220)	(0.9503)	(0.9645)	(0.9896)	(0.9940)

where $\chi_b^2(\beta)$ denotes the 100βth percentile of the chi-square random variable with b degrees of freedom, and an approximate $100(1-\alpha)\%$ confidence interval for $|\Sigma|$ is given by

$$(|\hat{\Sigma}|/c_{1-\alpha/2},\ |\hat{\Sigma}|/c_{\alpha/2}). \tag{5.12}$$

5.3. *Validity of the Approximations*

We evaluate the accuracies of the normal and chi-square approximations by the Monte Carlo method. For given α, N_i's, and p_i's, note that the 100αth percentile of $|\hat{\Sigma}|$ based on the normal approximation is given by

$$\exp(m + z_\alpha \sqrt{v}),$$

where m and v are as defined in Section 5.1, and the percentile based on the chi-square approximation is c_α given in (5.11). We estimated

$$P(|\hat{\Sigma}|/|\Sigma| \leqslant \exp(m + z_\alpha \sqrt{v})) \qquad \text{and} \qquad P(|\hat{\Sigma}|/|\Sigma| \leqslant c_\alpha)$$

using simulation consists of 100,000 runs. We used (5.3) to generate $|\hat{\Sigma}|/|\Sigma|$. For a good approximation, estimated proportion should be equal to the specified α. In Table III, we provide the approximate percentiles and the corresponding estimated cumulative probabilities (given in parentheses) for selected values of N_i's, p_i's, and α's. It is clear from the table values that the chi-square approximation is very satisfactory even for small samples whereas the normal approximation gives inaccurate percentiles even for large samples. Furthermore, examination of the table values indicate that the coverage probabilities of the confidence interval based on the normal distribution is noticeably smaller than the specified confidence levels. A similar comparison holds for the complete data case as well. Specifically, when there are no missing data, one should use Hoel's (1937) approximation to construct the confidence interval for the generalized variance. We also estimated percentiles for other dimensions and sample sizes configuration. Since they all exhibited a pattern similar to that in Table III, they are not reported here. Interested readers can find them in Hao (1999).

6. AN EXAMPLE

To illustrate the results of this article, we consider the following example. A sample of 20 observations, $Y_1, ..., Y_{20}$ is generated from a trivariate normal population with

$$\mu = \mathbf{0} \qquad \text{and} \qquad \Sigma = \begin{pmatrix} 7 & 1 & 3 \\ 1 & 5 & -1 \\ 3 & -1 & 2 \end{pmatrix}.$$

In order to create a monotone pattern data set, we discarded the last six observations on the third component. The simulated data are presented in

TABLE IV

The Simulated Data

5.4279	−1.1780	2.5194	0.8133	4.3632	−0.3055
2.8660	1.2840	0.1121	−4.0953	−0.2617	−1.3870
0.2184	0.5721	0.0069	3.9628	0.5064	2.8530
3.3805	1.6447	1.0339	−0.4451	−0.3697	1.0128
−3.2436	0.9633	−2.5652	1.3188	−0.1147	0.3009*
0.8957	−0.0033	−0.0024	3.3546	−3.8414	2.8740*
−3.4927	2.0351	−2.8056	1.9616	−0.4316	1.4052*
−4.0333	−1.8532	−0.9217	−0.4952	−2.7866	0.8983*
−1.2150	3.4360	−1.4132	1.0827	1.0524	−0.3738*
−1.6024	3.7717	−1.0030	2.8114	6.7408	−0.4069*

* These six entries were discarded to make a monotone sample.

Table IV. Thus, in our notations, we have $p_1 = 2$, $p_2 = 1$, $N_1 = 20$, and $N_2 = 14$.

6.1. *Hypothesis Testing about Σ*

The hypotheses are H_0: $\Sigma = \Sigma_0$ and H_a: $\Sigma \neq \Sigma_0$, where

$$\Sigma_0 = \begin{pmatrix} 8 & -2.5 & 3 \\ -2.5 & 4 & -1 \\ 3 & -1 & 2 \end{pmatrix}.$$

Note that Σ and Σ_0 are different only in the first 2×2 submatrix, and they are not a lot apart from each other. We chose this Σ_0 to check whether the proposed test is able to detect such a small difference between Σ and Σ_0.

Let T denote the lower triangular matrix with positive diagonals so that $\Sigma_0 = TT'$. For this example,

$$T^{-1} = \begin{pmatrix} T_{11} & 0 \\ T_{21} & T_{22} \end{pmatrix}^{-1} = \begin{pmatrix} 0.354 & 0 & 0 \\ 0.174 & 0.557 & 0 \\ -0.395 & 0.021 & 1.07 \end{pmatrix},$$

where T_{11} is of order 2×2. Make transformations $(x_{1j}, x_{2j}, x_{3j}) = (y_{1j}, y_{2j}, y_{3j})(T')^{-1}$ for $j = 1, ..., 14$, and $(x_{1j}, x_{2j}) = (y_{1j}, y_{2j})(T'_{11})^{-1}$ for $j = 15, ..., 20$ so that, under H_0, the vector x has a trivariate normal distribution with covariance matrix I_3.

For the transformed data, the sums of squares and the cross-products matrices are

$$V_{11,1} = \begin{pmatrix} 18.50 & 8.53 \\ 8.53 & 40.36 \end{pmatrix}$$

and

$$\begin{pmatrix} V_{11,2} & V_{12,2} \\ V_{21,2} & V_{22,2} \end{pmatrix} = \begin{pmatrix} 15.79 & 7.26 & 4.61 \\ 7.26 & 17.06 & -4.4 \\ 4.61 & -4.4 & 11.06 \end{pmatrix},$$

where $V_{11,2}$ is of order 2×2.

Furthermore, $a = 1.0253$, $b = 5.998$, $\Lambda_1 = 0.01113$, $\Lambda_2 = 0.5396$, $\mathrm{tr}(V_{21,2} V_{11,2}^{-1} V_{12,2}) = 4.48$, and $\Lambda = 6.3918 \times 10^{-4}$. The test statistic $-2 \ln \Lambda = 14.71$, with p-value $= P(a\chi_b^2 > 14.71) = 0.026$. The modified LRT statistic, based on partially complete data (the complete data obtained by discarding the additional data on the first $p_1 = 2$ components) is $-2\rho \ln \Lambda = 10.54$ with p-value 0.141 (see Muirhead, 1982, p. 359).

For the hypothesis

$$H_0: \mu = \mu_0, \Sigma = \Sigma_0 \quad \text{vs} \quad H_a: H_0 \text{ is not true},$$

let us suppose that $\mu_0' = (0.5, -0.5, 0)$ and Σ_0 is the same as above. Make the transformations on $(Y - \mu_0)$ using T. For the transformed data,

$$\bar{x}_{1,1}' = (-0.009, 0.707) \qquad \text{and} \qquad (\bar{x}_{1,2}', \bar{x}_{2,2}') = (-0.191, 0.778, 0.027).$$

Furthermore, $c = 1.017$, $d = 8.997$, $\Lambda = 6.3918 \times 10^{-4}$, $N_1 \bar{x}_{1,1}' \bar{x}_{1,1} + N_2 \bar{x}_{2,2}' \bar{x}_{2,2} = 10.006$, and $\delta = 4.2949 \times 10^{-6}$.

The statistic $-2 \ln \delta = 24.72$, with p-value $= 0.0054$. According to the standard procedure (see Anderson, 1984, p. 440) based on partially complete data with $N = 14$, the value of the statistic $-2 \ln \delta = 20.08$, with p-value $= 0.039$.

As this example illustrated, the modified LRT based on incomplete data provides more evidence against H_0 than the standard procedure based on partially complete data.

6.2. *Confidence Interval for* $|\Sigma|$

We compute a 95% confidence interval for $|\Sigma|$ using the data in Table IV. Recall that the mean $\mu = \mathbf{0}$, and the true value of $|\Sigma|$ is 10. Further, we have a monotone sample with $p_1 = 2$, $p_2 = 1$, $N_1 = 20$, and $N_2 = 14$. The sample statistics are,

$$S_{11,1} = \begin{pmatrix} 148.00 & -2.99 \\ & 117.31 \end{pmatrix},$$

and

$$\begin{pmatrix} S_{11,2} & S_{12,2} \\ & S_{22,2} \end{pmatrix} = \begin{pmatrix} 126.30 & -2.6177 & 58.82 \\ & 44.22 & -13.08 \\ & & 36.33 \end{pmatrix}.$$

We computed $|\hat{\Sigma}_{11,1}| = |\frac{1}{N_1} S_{11,1}| = 43.38$, and $|\hat{\Sigma}_{2.1}| = |\frac{1}{N_2} (S_{22,2} - S_{21,2} S_{12,2})| = 0.4105$. Hence, the MLE of $|\Sigma|$ is $|\hat{\Sigma}| = |\hat{\Sigma}_{11,1}| \cdot |\hat{\Sigma}_{2.1}| = 17.81$. A $100(1-\alpha)\%$ confidence interval for $|\Sigma|$ is

$$\left(\frac{N_1^2 N_2 \, |\hat{\Sigma}|}{\left(a\chi_b^2 \left(1 - \frac{\alpha}{2} \right) \right)^3}, \frac{N_1^2 N_2 \, |\hat{\Sigma}|}{\left(a\chi_b^2 \left(\frac{\alpha}{2} \right) \right)^2} \right),$$

where $a = 0.3443$ and $b = 43.20$. Using the table values $\chi_{43.2}^2(0.025) = 26.943$ and $\chi_{43.2}^2(0.975) = 63.23$, we compute the 95% confidence interval for $|\Sigma|$ as (9.67, 124.94), which contains the true value of $|\Sigma| = 10$.

ACKNOWLEDGMENTS

The authors are thankful to two reviewers and an editor for their valuable comments and suggestions.

REFERENCES

T. W. Anderson, Maximum likelihood estimates for a multivariate normal distribution when some observations are missing, *J. Amer. Statist. Assoc.* **52** (1957), 200–203.

T. W. Anderson, "An Introduction to Multivariate Statistical Analysis," Wiley, New York, 1984.

R. P. Bhargava, "Multivariate Tests of Hypotheses with Incomplete Data," Ph.D. dissertation, Stanford University, Stanford, CA, 1962.

S. Das Gupta, Properties of power functions of some tests concerning dispersion matrices of multivariate normal distributions, *Ann. Math. Statist.* **40** (1969), 697–701.

M. L. Eaton, "Some Problems in Covariance Estimation," Technical Report 49, Department of Statistics, Stanford University, Stanford, CA, 1970.

J. Hao, "Inferences for the Normal Covariance Matrix with Incomplete Data," Ph.D. dissertation, Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA, 1999.

P. G. Hoel, A significance test for component analysis, *Ann. Math. Statist.* **8** (1937), 149–158.

K. G. Jinadasa and D. S. Tracy, Maximum likelihood estimation for multivariate normal distribution with monotone sample, *Comm. Statist. A*, *Theory-Meth.* **21** (1992), 41–50.

K. Krishnamoorthy and M. Pannala, Some simple test procedures for normal mean vector with incomplete data, *Ann. Inst. Statist. Math.* **50** (1998), 531–542.

K. Krishnamoorthy and M. Pannala, Confidence estimation of a normal mean vector with incomplete data, *Canad. J. Statist.* **27** (1999), 395–407.

K. H. Li, T. E. Raghunathan, and D. B. Rubin, Large-sample significance levels from multiply imputed data using moment-based statistics and an F reference distribution, *J. Amer. Statist. Assoc.* **86** (1991), 1065–1073.

R. J. A. Little and D. B. Rubin, "Statistical Analysis with Missing Data," Wiley, New York, 1987.

F. M. Lord, Estimation of parameters from incomplete data, *J. Amer. Statist. Assoc.* **50** (1955), 870–876.

J. S. Mehta and J. Gurland, A test of equality of means in the presence of correlation and missing values, *Biometrika* **60** (1969), 211–213.

X. L. Meng, Multiple-imputation inferences with uncongenial sources of input, *Statist. Sci.* **9** (1994), 538–573.

D. F. Morrison, A test for equality of means of correlated variates with missing data on one response, *Biometrika* **60** (1973), 101–105.

U. D. Naik, On testing equality of means of correlated variables with incomplete data, *Biometrika* **62** (1976), 615–622.

R. J. Muirhead, "Aspects of Multivariate Statistical Theory," Wiley, New York, 1982.

D. B. Rubin, Inference and missing data (with Discussion), *Biometrika* **63** (1976), 581–592.

D. B. Rubin and N. Schenker, Multiple imputation for interval estimation from simple random samples with ignorable nonresponse, *J. Amer. Statist. Assoc.* **81** (1986), 366–374.

F. E. Satterthwaite, An approximate distribution of estimates of variance components, *Biometrika* **33** (1946), 110–114.

D. Sharma and K. Krishnamoorthy, Improved minimax estimators of normal covariance and precision matrices from incomplete samples, *Calcutta Statist. Assoc. Bull.* **34** (1985), 23–42.

M. Siotani, T. Hayakawa, and Y. Fujikoshi, "Modern Multivariate Statistical Analysis: A Graduate Course and Handbook," American Sciences Press, Columbus, Ohio, 1985.